



# Plant phylogenomics based on genome-partitioning strategies: Progress and prospects



Xiangqin Yu <sup>a</sup>, Dan Yang <sup>a, c</sup>, Cen Guo <sup>b, c</sup>, Lianming Gao <sup>a, \*</sup>

<sup>a</sup> Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan, 650201, China

<sup>b</sup> Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan, 650201, China

<sup>c</sup> Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan, 650201, China

## ARTICLE INFO

### Article history:

Received 24 April 2018

Received in revised form

26 June 2018

Accepted 27 June 2018

Available online 30 June 2018

(Editor: Zhekun Zhou)

### Keywords:

Plant phylogenomics

Next-generation sequencing

Whole-genome sequencing

Genome skimming

RAD-Seq

Targeted capture

## ABSTRACT

The rapid expansion of next-generation sequencing (NGS) has generated a powerful array of approaches to address fundamental questions in biology. Several genome-partitioning strategies to sequence selected subsets of the genome have emerged in the fields of phylogenomics and evolutionary genomics. In this review, we summarize the applications, advantages and limitations of four NGS-based genome-partitioning approaches in plant phylogenomics: genome skimming, transcriptome sequencing (RNA-seq), restriction site associated DNA sequencing (RAD-Seq), and targeted capture (Hyb-seq). Of these four genome-partitioning approaches, targeted capture (especially Hyb-seq) shows the greatest promise for plant phylogenetics over the next few years. This review will aid researchers in their selection of appropriate genome-partitioning approaches to address questions of evolutionary scale, where we anticipate continued development and expansion of whole-genome sequencing strategies in the fields of plant phylogenomics and evolutionary biology research.

Copyright © 2018 Kunming Institute of Botany, Chinese Academy of Sciences. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

The invention of DNA sequencing was one of the most transformative events in biology (Sanger et al., 1977). Two decades later, the first genome sequence of a bacterium (*Haemophilus influenzae*) was reported (Fleischmann et al., 1995), quickly followed by several other organisms (The *C. elegans* Sequencing Consortium, 1998; Goffeau et al., 1996) culminating in The Human Genome Project (HGP), which was completed in 2004 (International Human Genome Sequencing Consortium, 2004). Thereafter, the whole genomes of a large number of organisms were sequenced followed by the rapid development of next-generation sequencing (NGS) technology, which substantially decreased sequencing costs. As such, the rapid expansion of NGS provided a powerful array of tools to address fundamental biological questions at multiple scales (Jones and Good, 2016), marking the genomic era in biological research.

Phylogenomics is a new and exciting synthesized discipline (Delsuc et al., 2005; Eisen, 1998) that is at the intersection of evolution and genomics (Eisen and Fraser, 2003). The main goal of phylogenomics is to infer species relationships using genomic data, as well as to gain knowledge of the mechanisms of molecular evolution based on the evolutionary history of species (Philippe et al., 2005). Understanding phylogenetic relationships between organisms is a prerequisite of almost all evolutionary studies (Delsuc et al., 2005; Zhang et al., 2012), although many plant phylogenies based on traditional DNA-fragments remain unresolved at all evolutionary scales due to a lack of informative sites. Phylogenomics can potentially resolve species relationships by making use of vast sequence data as well as gene order, insertions and deletions (indels), retroposon integrations, and gene fusion and fission events (Rokas and Holland, 2000). Access to genomic data could also potentially alleviate previous problems of phylogenetics that resulted from stochastic error (limitation of sampling few genes) by expanding the number of characters (Delsuc et al., 2005). Thus, phylogenomics provides a window for better understanding the evolutionary relationships of plants using genome-scale data and generating a more robust picture of the next generation Tree of Life.

\* Corresponding author. Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, 132# Lanhei Road, Kunming, 650201, Yunnan, China. Fax: +86 0871 65225286.

E-mail address: [gaolm@mail.kib.ac.cn](mailto:gaolm@mail.kib.ac.cn) (L. Gao).

Peer review under responsibility of Editorial Office of Plant Diversity.

From first generation (Sanger) DNA sequencing to the second (massively parallel) and third (real-time, single-molecule) generation DNA sequencing, gathering genomic data has become much more convenient and cost-effective (Shendure et al., 2017). Among the three genomes of plants, the plastid genome evolves more rapidly and has lower inter- and intramolecular recombination rates than the mitochondrial genome (Lonsdale et al., 1988; Palmer and Herbon, 1988). Furthermore, the plastid genome can be more easily sequenced than the nuclear genome. Despite earlier debate on genome-based phylogenies, Martin et al. (2005) argued for the critical role that the plastid genome plays in plant phylogenetics. Sequencing the plastid genome was the most common approach in the early stages of plant phylogenetic and evolution research, providing an efficient method for investigating the evolutionary relationships and basal lineages of angiosperms (Goremykin et al., 2003, 2004; Jansen et al., 2007; Moore et al., 2007).

Since 2015, improved NGS technologies, lower genome sampling costs, and the development of statistical methods have greatly expanded the use of genomic data in plant phylogenetics (Barrett et al., 2016; Gao et al., 2010). Various genome-partitioning strategies to sequence selected subsets of the genome, such as genome skimming, transcriptome sequencing (RNA-seq), restriction site associated DNA sequencing (RAD-Seq), and targeted capture have emerged as powerful alternatives to whole-genome sequencing (WGS) in ecological and evolutionary genomic studies (Jones and Good, 2016) as well as studies in plant phylogenetics. For example, as of April 2018, more than 2800 records of whole plastid genomes of plants have been deposited in GenBank. In most cases, sequences were obtained by long-range polymerase chain reaction (PCR) (Yang et al., 2014) and genome skimming (Straub et al., 2012) with NGS. Based on such data, plastid phylogenomics has undoubtedly been serving as an effective approach for uncovering deep-level relationships of intractable and even rapidly radiating plant groups (Barrett et al., 2013; Ma et al., 2014; Ross et al., 2015; Wysocki et al., 2015). In addition, nuclear and mitochondrial genomes are expected to play an increasing role in plant phylogenetics in the future (Liu et al., 2014; Vargas et al., 2017; Zeng et al., 2014; Zhang et al., 2012). The international multidisciplinary consortium “1000 Plants (1KP) project” aims to generate transcriptome data from over 1300 green plants, including all of the major lineages across the Viridiplantae clade, underscoring the considerable effort this team has made to investigate the value of nuclear genomic data for plant phylogenomic analyses (Matasci et al., 2014). Furthermore, making use of the whole nuclear genome or data from resequencing for plant phylogenomics and population genomics is also at the forefront of many studies (Sollars et al., 2016; Teh et al., 2017; Zhang et al., 2017a). We have now reached the point where these approaches to answering fundamental evolutionary questions have transformed our research to an unprecedented degree.

## 2. The application of genomic data in plant phylogenetics

Over the past few years, NGS-based genomic data have made a profound impact on phylogenetics (Jarvis et al., 2014; Misof et al., 2014). Several genome-partitioning strategies to sequence selected subsets of the genome, including genome skimming, RNA-seq, RAD-Seq, and targeted capture, have emerged as powerful tools in plant phylogenomics. Here we summarize the applications, advantages, and limitations of four NGS-based genome-partitioning approaches in plant phylogenomics (Table 1). We hope this review will help researchers choose the appropriate approaches to address phylogenetic questions at various evolutionary scales.

### 2.1. Phylogenomics using plastid genomes

Land plant plastid genomes share the typical quadripartite structure including two rRNA-containing inverted repeats (IRs) and two unequal single-copy regions (Raubeson and Jansen, 2005). The plastid genome size of land plants ranges from 11 kb to 217 kb (Bellot and Renner, 2015; Guisinger et al., 2010). Even though some rearrangements occur in certain lineages, the structure of land plant plastid genomes is generally conserved (Gao et al., 2010). The first two plastid genomes were determined by constructing a set of overlapping restriction endonuclease fragments and Sanger sequencing (Ohya et al., 1986; Shinozaki et al., 1986). Faster and more cost-effective approaches were then developed, including 1) shearing, cloning and sequencing the isolated pure cpDNA; 2) amplification using long PCR; and 3) construction of bacterial artificial chromosome (BAC) or Fosmid libraries (Jansen et al., 2005). Using conserved primers based on available plastid genomes, long range PCR was an effective tool for obtaining whole genomes for phylogenetic analyses such as identifying basal angiosperm lineages (Goremykin et al., 2003, 2004). However, these studies mostly rely on traditional Sanger sequencing, a process that was time-consuming and expensive. With the advent of NGS in 2005, followed by the development of library-construction-based NGS platforms (such as Roche 454, Solexa, SOLiD and Helicos), sequencing entered a new era characterized by high-throughput and cost-efficiency (Shendure et al., 2017). Angiosperm phylogenies were then greatly improved by including more genes from plastid genomes and extending taxa sampling, mainly through sequencing whole plastid genomes (Moore et al. 2007, 2010). By integrating more plastid genomes as templates, Yang et al. (2014) and Zhang et al. (2016) reported nine and fifteen novel universal primer pairs for amplification of whole plastid genomes of angiosperms, respectively. This approach was subsequently used in phylogenomic analyses of several plant lineages such as Theaceae, Rosaceae and Cornales (Fu et al., 2017; Yu et al., 2017; Zhang et al., 2017b). The disadvantage of this approach is the need of high quality genomic DNA isolated from fresh material or high quality material quickly dried in silica-gel after collection.

Genome skimming was proposed as a way of ‘navigating the tip of the genomic iceberg’ (Nock et al., 2011). This approach consists of shallow sequencing of genomic DNA that results in comparatively deep sequencing of the high-copy fraction of the genome (Straub et al., 2012). Genome skimming is an efficient approach for obtaining the complete plastid genome (ptDNA), a large fraction of the mitochondrial genome (mtDNA) and the nuclear ribosomal cluster (nrDNA). It is cost-effective, and can tolerate low-quality samples (fresh, silica-gel-preserved leaves or even Herbarium specimens up to 146 years old) (Bakker et al., 2015; Dodsworth, 2015). As such, the use of genome skimming has contributed to numerous advances in our understanding of species relationships across a broad phylogenetic range of taxa. For example, shotgun-sequencing-based genome skimming of the pantropical tree family Chrysobalanaceae yielded more robust phylogenetic relationships than previous studies (Male et al., 2014). Barrett et al. (2016) obtained 39 plastomes using genome skimming to investigate deep phylogenetic relationships and extensive rate variation among palms and other commelinid monocots. The pattern of reticulate evolution in a species-rich and recently diverged Andean genus *Diplostephium* was revealed by integrating phylogenetic signal of genomic regions with different inheritance patterns using genome skimming and ddRADseq (Vargas et al., 2017).

Organellar genomes are generally inherited uniparentally; only 14% of angiosperms inherit plastids biparentally (Corriveau and Coleman, 1988). Accordingly, Gitzendanner et al. (2018) recently

**Table 1**  
Comparison of the four commonly used genome-partitioning approaches.

	Genome skimming	RNA-seq	RAD-seq	Hyb-seq
Demand for plant materials	Fresh, silica-gel dried plant tissues and specimen	Fresh plant tissues	Fresh, silica-gel dried plant tissues	Fresh, silica-gel dried plant tissues and specimen
Demand for DNA template quality	Low	High	Medium	Low
Applicable to specimen	Yes	No	No	Yes
Material for sequencing	Total genomic DNA	cDNA	Restriction fragments	Captured loci using probes
Genome data obtained	Complete plastid genome, nrDNA, partial mitochondrial genome, coding and non-coding genes	Randomly sequenced loci of vast majority of nuclear genome; coding genes	Loci with single nucleotide polymorphism (SNP) mainly from nuclear genome; coding and non-coding genes	Targeted nuclear, plastid and/or mitochondrial loci; coding and non-coding genes
Targeted loci sequenced	Yes	No	No	Yes
Identification of orthologs	Easy	Relatively easy	Difficult	Easy
Missing data among species	No	Yes	Yes	No
Taxonomic levels for phylogenetic relationships	All levels from shallow to deep	Deep levels, above intra-generic	Shallow levels, below inter-generic	All levels from shallow to deep, above intraspecific

suggested that the plastid genome provides only one perspective on plant evolutionary history; thus, a plastid-based tree should not be blindly accepted as the backbone tree for Viridiplantae. Organelle capture–introgression of the organellar genomes from one species into another may cause phylogenetic inconsistencies between organellar and nuclear trees (Huang et al., 2014; Stegemann et al., 2012; Tsitroni et al., 2003; Yi et al., 2015). Recombination and gene conversion that have occurred in the plastid genome might also introduce biases and errors to phylogenetic reconstruction (Davis et al., 2014). Sullivan et al. (2017) even reported the extreme case of interspecific plastome recombination in *Picea* (Pinaceae), which might result in discordant plastid phylogenies. In contrast, nuclear genes are biparentally inherited and contain abundant genetic information, thus more data from nuclear genomes are needed to provide alternative evidence for phylogenetic relationship reconstructions (Lee et al., 2011; Zimmer, 2013).

## 2.2. Phylogenomics using transcriptome sequencing

Comparative genomic approaches among model species have been successfully used to identify a large set of orthologous nuclear markers that can now be used in phylogenetic studies, opening new avenues for molecular systematics (Duarte et al., 2010; Soltis et al., 2013). RNA-seq of two or three species per clade of the studied taxa produces a large number of single-copy candidate loci that can be screened for substitution rate, ease of amplification, and position of introns in other taxa (Harrison and Kidner, 2011). Alternatively, transcriptomes can directly be used for phylogenetic analyses. Zhang et al. (2012) used genome comparisons between seven angiosperms and one moss species to identify 1083 highly conserved, low-copy nuclear genes, five of which were valuable in reconstructing a highly resolved angiosperm phylogeny. While the angiosperm phylogeny that was reconstructed using these five genes was largely congruent with phylogenies previously inferred from organellar genes, several new placements were uncovered for some lineages. Integrating 26 new transcriptomes with previously reported orthologous genes, 59 carefully selected low-copy nuclear genes were used to build a highly supported deep-level (among eight clades) angiosperm phylogeny. Molecular clock estimates of mesangiospermae diversification have been used to illuminate a possible link between origins of some insects and the early angiosperm radiation (Zeng et al., 2014). Furthermore, orthologous nuclear genes derived from transcriptomes were used to uncover robust phylogenies for eudicot, Caryophyllales, and several species-rich families (Rosaceae, Brassicaceae, and Asteraceae) (Huang et al., 2015, 2016; Xiang et al., 2016; Yang et al., 2015; Zeng et al., 2017b).

The 1000 Plant Genomes Project (1KP), consisting of transcriptomes from over 1300 species representing the diversity of

green plants, is the first international collaboration on a large-scale transcriptome sequencing project for plants, and has provided evidence for the resolution of phylogenetic uncertainties (Granados Mendoza et al., 2015; Matasci et al., 2014). For instance, the origin and early diversification of land plants has also been investigated through phylotranscriptomic (including 1KP data) analysis of 852 nuclear genes and 1,701,170 aligned sites (Wickett et al., 2014). However, transcriptomics requires living tissue for RNA extraction, and thus many existing tissue collections are unusable (Soltis et al., 2013). Therefore, a substantial amount of effort in future studies will be placed on resampling. In addition, RNA should be sampled from the same types of tissue and from individuals at the same life-history stage to obtain as many orthologous loci across samples as possible (Lemmon and Lemmon, 2013).

## 2.3. Phylogenomics using restriction site-associated DNA sequencing

Restriction site-associated DNA sequencing (RAD-Seq) and its related methods rely on the conservation of enzyme recognition sites (e.g., GBS: Genotyping-By-Sequencing; SLAF-seq: Specific-Locus Amplified Fragment Sequencing; SBG: Sequence-Based Genotyping). This technique can be used to survey hundreds or thousands of unlinked genetic markers adjacent to restriction sites from the nuclear genome (Baird et al., 2008; Elshire et al., 2011; Peterson et al., 2012). This approach mainly involves digestion of genomic DNA samples with restriction enzymes, size selection of a subset of the restriction fragments, PCR amplification and high-throughput sequencing of the size-selected fragments (Andrews et al., 2016). Typically, this technique has been used for rapid single nucleotide polymorphism (SNP) discovery and genotyping for genetic mapping of large populations in a variety of organisms (Baird et al., 2008; Baxter et al., 2011; Elshire et al., 2011; Pfender et al., 2011).

Since RAD-seq possesses several facilitative advantages, it has shown great promise and been broadly used to resolve phylogenetic relationships (Baird et al., 2008; Baxter et al., 2011; Henning et al., 2014; Pfender et al., 2011). First, RAD-seq can create a reduced representation of the genome, allowing detection of numerous informative SNPs. These markers provide unprecedented resolution of the framework phylogenies for several complex biological taxa, especially at intergeneric, interspecific and intraspecific taxonomic levels, such as Arundinarieae of Graminaeae (Wang et al., 2017), American oak clade (Hipp et al., 2014, 2018), *Carex* (Escudero et al., 2014; Massatti et al., 2016), *Pedicularis* section *Cyathophora* (Eaton and Ree, 2013), and *Primula tibetica* (Ren et al., 2017). Second, this approach serves as a powerful tool that has largely reduced the limitations of phylogenetic

reconstruction for non-genome species (Cariou et al., 2013; Etter and Johnson, 2012). Third, this high-throughput approach is relatively cost- and time-effective compared to other approaches such as whole genome sequencing (WGS) (Yang et al., 2016).

Although RAD-seq provides a potential approach to resolve shallow evolutionary timescales and phylogenetic relationships, its utility in resolving deep phylogenetic relationships is limited (Jones and Good, 2016). RAD-seq is mostly used for phylogenetic inference below the genus level (e.g., Eaton and Ree, 2013; Eaton et al., 2017; Emerson et al., 2010). However, RAD-seq data seem to be unreliable beyond moderate phylogenetic inference because restriction cut sites are usually not conserved across distinct taxa, leading to high levels of missing data and incorrect topologies (Leaché et al., 2015; Rubin et al., 2012; Wagner et al., 2013). In addition, RAD-seq methods introduce several unique potential sources of error and bias in RAD-seq methods when doing phylogenetic studies. First, selecting a fragment size range by manual excision is subject to some human error, potentially leading to fewer orthologous fragments across individuals (McCormack et al., 2013b). Second, because the RAD tags are short (<300bp) and not targeted (Andrews et al., 2016), coverage can be difficult to estimate and the identification of orthologous fragments may vary among different assembly methods (Wang et al., 2017). Third, RAD-seq data notably suffer from a potentially large amount of missing data, which has triggered a serious debate on the effect of such missing data in reconstructing phylogenetic relationships (Eaton et al., 2017; Huang and Knowles, 2014; Leaché et al., 2015; Wagner et al., 2013). Another potential drawback is that most methods based on restriction digests are geared toward SNP generation, which may not be ideal for phylogenetic construction (Jones and Good, 2016).

#### 2.4. Phylogenomics using targeted-capture

Targeted capture (targeted enrichment, hybridization enrichment, Hyb-Seq: combination of target enrichment and genome skimming approaches) is a technique that uses a hybridization reaction involving custom-designed short RNA or DNA probes in solution or on an array to capture thousands of target loci with sequences similar to the set of probes from fragmented genomic DNA libraries (Nicholls et al., 2015; Senapathy et al., 2010). Target capture, which allows simultaneous capture of low-copy nuclear genes and high-copy genomic targets (Weitemier et al., 2014), alleviates the limitations associated with RNA-seq and provides a prospective approach for plant phylogenetics. This high-throughput approach is more cost-effective at obtaining large data sets of orthologous loci across many individuals compared to WGS and multiplex PCR (Olson, 2007). A major challenge, however, is identifying the genomic sequences to be used for capture probes designed in non-reference species where a priori knowledge of target sequences is required (Elshire et al., 2011; Nicholls et al., 2015). This hurdle, nevertheless, can be resolved by utilizing de novo genomes for a large number of species that have already sequenced and the recent rapid accumulation of various kinds of genomic data (such as transcriptomes, genome skimming and RAD-seq). Such approaches have been successfully applied to resolve phylogenetic relationships of plants as well as other organisms (Brandley et al., 2015; Eytan et al., 2015; Fragoso-Martínez et al., 2017; Weitemier et al., 2014).

Targeted capture is appealing because it can theoretically be applied to any species with a de novo sequence assembly, (e.g., de novo whole genome sequencing, de novo RNA-seq transcriptomes or expressed sequence tag (EST) data), which have been or could be generated in the future (Bi et al., 2012, 2013). The flexibility of targeted capture for phylogenomic studies offers a tremendous advantage over other methods (Jones and Good, 2016). Customized

capture designs from a set of probes with NGS can target 1) hundreds or thousands of orthologous loci (Faircloth et al., 2012; Mandel et al., 2014; Valderrama et al., 2018); 2) slowly or quickly evolving loci, including coding genes and non-coding regions (Lemmon and Lemmon, 2013; McCormack et al., 2013a); and 3) nuclear or organelle loci (Hedtke et al., 2013; Ilves and Lopez-Fernandez, 2014). Furthermore, targeted capture can work on a range of genomic DNA quality from fresh, silica-gel dried materials, even herbarium specimens of plants (Hart et al., 2016), whereas RNA-seq transcriptomes require fresh materials and RAD-seq requires high quality DNA isolated from fresh and silica-gel dried leave of plants.

Selecting loci with appropriate evolutionary rates is very important when resolving a given relationship (Philippe et al., 2011). For example, deep phylogenetic nodes are resolved by using slowly evolving loci that retain signals of orthology across distant taxa (McCormack et al., 2012; Schott et al., 2017). High-throughput targeted capture of slowly evolving ultraconserved element (UCE) markers has been used to resolve deep nodes in the phylogenies of vertebrates (Crawford et al., 2015; Faircloth et al., 2012) as well as complex phylogenetic histories in flowering plants (Folk et al., 2017; Mandel et al., 2014). Capture of UCES is appropriate for resolving deep node relationships because a core set of loci can be identified from highly divergent reference genomes and can then be widely used for diverse taxa without continually redesigning custom probes (Jones and Good, 2016). In contrast, protein-coding sequences may be more suitable for plant phylogenetic reconstruction at moderate-to-deep evolutionary scales; these sequences are less conserved than UCES, but more conserved than non-coding sequences (Mandel et al., 2014; Nicholls et al., 2015; Valderrama et al., 2018). To maximize phylogenetically informative sites at shallow evolutionary timescales (the inter-generic or intra-generic tips of the phylogeny), non-coding regions should be combined with coding regions, especially for recently radiated taxa. Phylogenetic studies based on target enrichment benefit from using a large number of putatively independent nuclear loci and their combination with plastid and mitochondrial genomes (Schmickl et al., 2016).

Targeted capture has emerged as a powerful approach in the genomic era to address plant phylogenetic questions and will be widely used for phylogenomics studies at various evolutionary scales. Hyb-Seq, a modified technique that combines target enrichment and genome skimming (Weitemier et al., 2014), has the advantages of both target enrichment and genome skimming, and will exhibit great promise for plant phylogenomics in the next few years.

### 3. Prospects in plant phylogenomics

Nuclear genomic data have been increasingly explored for phylogenomics. For example, nuclear repeat regions were used in *Solanum* section *Lycopersicon* to test the usefulness and power of phylogenomic analyses at inter- and intraspecific levels (Dodsworth et al., 2016). This approach provides additional evidence that might complement the results from organellar and the nuclear ribosomal cistron obtained from genome skimming. Currently, obtaining high-quality, low-cost genome sequences for a taxon of interest is routine. Third generation sequencing – real-time, single-molecule sequencing (PacBio and nanopore sequencing) – will be able to generate reads over 10 kb (or even 100 kb) (Bayley, 2015; Deamer et al., 2016; Eid et al., 2009; Shendure et al., 2017). Genome sequencing will benefit from the innovation of these sequencing technologies. Whole genome sequences have been used to infer the evolutionary relationships between closely related species of a flycatcher species complex

(*Ficedula*, Muscipidae) and the early branches in the tree of life of modern birds and mammals (Jarvis et al., 2014; Nater et al., 2015; Sims et al., 2009). Resequencing, or mapping sequence reads to a reference genome to identify genetic variants, is less time-consuming than genome assembly (Shendure et al., 2017). It has been used to infer the geographic origin and migration history of the brown rat (*Rattus norvegicus*) (Zeng et al., 2017a), the genetic diversity of European ash trees (*Fraxinus excelsior*) (Sollars et al., 2016), and the history of apple domestication along the Silk Road (Duan et al., 2017). These achievements represent a prospective picture for resolving intractable plant relationships using whole genome and genome resequencing. Notably, the China National GeneBank and BGI-Shenzhen will lead the 10KP (10,000 Plants) Genome Sequencing Project, which will sequence and characterize representative genomes from every major clade of embryophyte, green algae, and protist (excluding fungi) within the next five years (Cheng et al., 2018). Once the project is finished, it will provide valuable genome resources for addressing numerous fundamental questions in evolutionary and comparative genomics and will greatly benefit our understanding of plant evolution and diversity. With the improvement and innovation of statistical and computational abilities for genome-sequence-based phylogeny (Mirarab et al., 2014; Stamatakis et al., 2012), phylogenomics embraces a new era of discovery based on large amounts of genomic data and powerful analytical methods. The merger of phylogenomics with other biological disciplines, such as biogeography and ecology, will greatly advance our understanding of the origins and evolution of earth's biodiversity.

## Acknowledgements

We are grateful to Profs De-Zhu Li, Jun-Bo Yang and Hong-Tao Li for their discussion of the early version of the manuscript. We also thank Prof. Kevin S. Burgess from Columbus State University for improving the English of the manuscript. This study was supported by the Large-scale Scientific Facilities of the Chinese Academy of Sciences (Grant No: 2017-LSFGBOWS-01), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB31000000), and the Program of Science and Technology Talents Training of Yunnan Province (2017HA014).

## References

- Andrews, K.R., Good, J.M., Miller, M.R., et al., 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.* 17, 81.
- Baird, N.A., Etter, P.D., Atwood, T.S., et al., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3, e3376.
- Bakker, F.T., Lei, D., Yu, J., et al., 2015. Herbarium genomics: plastome sequence assembly from a range of herbarium specimens using an Iterative Organelle Genome Assembly pipeline. *Biol. J. Linn. Soc.*
- Barrett, C.F., Davis, J.L., Leebens-Mack, J., et al., 2013. Plastid genomes and deep relationships among the commelinid monocot angiosperms. *Cladistics* 29, 65–87.
- Barrett, C.F., Baker, W.J., Comer, J.R., et al., 2016. Plastid genomes reveal support for deep phylogenetic relationships and extensive rate variation among palms and other commelinid monocots. *New Phytol.* 209, 855–870.
- Baxter, S.W., Davey, J.W., Johnston, J.S., et al., 2011. Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. *PLoS One* 6, e19315.
- Bayley, H., 2015. Nanopore sequencing: from imagination to reality. *Clin. Chem.* 61, 25–31.
- Bellot, S., Renner, S.S., 2015. The plastomes of two species in the endoparasite genus *Pilosyles* (Apodanthaceae) each retain just five or six possibly functional genes. *Genome Biol. Evol.* 8, 189–201.
- Bi, K., Vanderpool, D., Singhal, S., et al., 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genom.* 13.
- Bi, K., Linderoth, T., Vanderpool, D., et al., 2013. Unlocking the vault: next-generation museum population genomics. *Mol. Ecol.* 22, 6018–6032.
- Brandley, M.C., Bragg, J.G., Singhal, S., et al., 2015. Evaluating the performance of anchored hybrid enrichment at the tips of the tree of life: a phylogenetic analysis of Australian *Eugongylus* group scincid lizards. *BMC Evol. Biol.* 15.
- Cariou, M., Duret, L., Charlat, S., 2013. Is RAD-seq suitable for phylogenetic inference? An in silico assessment and optimization. *Ecol. Evol.* 3, 846–852.
- Cheng, S., Melkonian, M., Smith, S.A., et al., 2018. 10KP: a phylodiverse genome sequencing plan. *GigaScience* 7, 4880447.
- Corriveau, J.L., Coleman, A.W., 1988. Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 Angiosperm species. *Am. J. Bot.* 75, 1443–1458.
- Crawford, N.G., Parham, J.F., Sellas, A.B., et al., 2015. A phylogenomic analysis of turtles. *Mol. Phylogenet. Evol.* 83, 250–257.
- Davis, C.C., Xi, Z., Mathews, S., 2014. Plastid phylogenomics and green plant phylogeny: almost full circle but not quite there. *BMC Biol.* 12, 11.
- Deamer, D., Akeson, M., Branton, D., 2016. Three decades of nanopore sequencing. *Nat. Biotechnol.* 34, 518–524.
- Delsuc, F., Brinkmann, H., Philippe, H., 2005. Phylogenomics and the reconstruction of the tree of life. *Nature* 6, 361–375.
- Dodsworth, S., 2015. Genome skimming for next-generation biodiversity analysis. *Trends Plant Sci.* 20, 525–527.
- Dodsworth, S., Chase, M.W., Sarkinen, T., et al., 2016. Using genomic repeats for phylogenomics: a case study in wild tomatoes (*Solanum* section *Lycopersicon*: solanaceae). *Biol. J. Linn. Soc.* 117, 96–105.
- Duan, N., Bai, Y., Sun, H., et al., 2017. Genome re-sequencing reveals the history of apple and supports a two-stage model for fruit enlargement. *Nat. Commun.* 8, 249.
- Duarte, J.M., Wall, P.K., Edger, P.P., et al., 2010. Identification of shared single copy nuclear genes in *Arabidopsis*, *Populus*, *Vitis* and *Oryza* and their phylogenetic utility across various taxonomic levels. *BMC Evol. Biol.* 10, 61.
- Eaton, D.A., Ree, R.H., 2013. Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: orobanchaceae). *Syst. Biol.* 62, 689–706.
- Eaton, D.A., Spriggs, E.L., Park, B., et al., 2017. Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. *Syst. Biol.* 66, 399–412.
- Eid, J., Fehr, A., Gray, J., et al., 2009. Real-time DNA sequencing from single polymerase molecules. *Science* 323, 133–138.
- Eisen, J.A., 1998. Phylogenomics: improving functional predictions for uncharacterized genes by evolutionary analysis. *Genome Res.* 8, 163–167.
- Eisen, J.A., Fraser, C.M., 2003. Phylogenomics: intersection of evolution and genomics. *Science* 300, 1706–1707.
- Elshire, R.J., Glaubitz, J.C., Sun, Q., et al., 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6, e19379.
- Emerson, K.J., Merz, C.R., Catchen, J.M., et al., 2010. Resolving postglacial phylogeography using high-throughput sequencing. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16196–16200.
- Escudero, M., Eaton, D.A., Hahn, M., et al., 2014. Genotyping-by-sequencing as a tool to infer phylogeny and ancestral hybridization: a case study in *Carex* (Cyperaceae). *Mol. Phylogenet. Evol.* 79, 359–367.
- Etter, P.D., Johnson, E., 2012. RAD paired-end sequencing for local de novo assembly and SNP discovery in non-model organisms. *Data Production and Analysis in Population Genomics*. Humana Press, Totowa, NJ, pp. 135–151.
- Eytan, R.I., Evans, B.R., Dornburg, A., et al., 2015. Are 100 enough? Inferring acanthomorph teleost phylogeny using Anchored Hybrid Enrichment. *BMC Evol. Biol.* 15.
- Faircloth, B.C., McCormack, J.E., Crawford, N.G., et al., 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary time-scales. *Syst. Biol.* 61, 717–726.
- Fleischmann, R.D., Adams, M.D., White, O., et al., 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269, 496–512.
- Folk, R.A., Mandel, J.R., Freudenstein, J.V., 2017. Ancestral gene flow and parallel organellar genome capture result in extreme phylogenomic discord in a lineage of angiosperms. *Syst. Biol.* 66, 320–337.
- Fragoso-Martínez, I., Salazar, G.A., Martínez-Gordillo, M., et al., 2017. A pilot study applying the plant anchored hybrid enrichment method to New World sages (*Salvia* subgenus *Calospatha*; Lamiaceae). *Mol. Phylogenet. Evol.* 117, 124–134.
- Fu, C.N., Li, H.T., Milne, R., et al., 2017. Comparative analyses of plastid genomes from fourteen Cornales species: inferences for phylogenetic relationships and genome evolution. *BMC Genom.* 18, 956.
- Gao, L., Su, Y.J., Wang, T., 2010. Plastid genome sequencing, comparative genomics, and phylogenomics: current status and prospects. *J. Systemat. Evol.* 48, 77–93.
- Gitzendanner, M.A., Soltis, P.S., Wong, G.K.S., et al., 2018. Plastid phylogenomic analysis of green plants: a billion years of evolutionary history. *Am. J. Bot.* 105, 291–301.
- Goffeau, A., Barrell, B.G., Bussey, H., et al., 1996. Life with 6000 genes. *Science* 274, 546–567.
- Goremykin, V.V., Hirsch-Ernst, K.I., Wölfl, S., et al., 2003. Analysis of the *Amborella trichopoda* chloroplast genome sequence suggests that *Amborella* is not a basal angiosperm. *Mol. Biol. Evol.* 20, 1499–1505.
- Goremykin, V.V., Hirsch-Ernst, K.I., Wölfl, S., et al., 2004. The chloroplast genome of *Nymphaea alba*: whole-genome analyses and the problem of identifying the most basal angiosperm. *Mol. Biol. Evol.* 21, 1445–1454.
- Granados Mendoza, C., Naumann, J., Samain, M.S., et al., 2015. A genome-scale mining strategy for recovering novel rapidly-evolving nuclear single-copy

- genes for addressing shallow-scale phylogenetics in *Hydrangea*. *BMC Evol. Biol.* 15.
- Guisinger, M.M., Kuehl, J.V., Boore, J.L., et al., 2010. Extreme reconfiguration of plastid genomes in the angiosperm family Geraniaceae: rearrangements, repeats, and codon usage. *Mol. Biol. Evol.* 28, 583–600.
- Harrison, N., Kidner, C.A., 2011. Next-generation sequencing and systematics: what can a billion base pairs of DNA sequence data do for you? *Taxon* 60, 1552–1566.
- Hart, M.L., Forrest, L.L., Nicholls, J.A., et al., 2016. Retrieval of hundreds of nuclear loci from herbarium specimens. *Taxon* 65, 1081–1092.
- Hedtke, S.M., Morgan, M.J., Cannatella, D.C., et al., 2013. Targeted enrichment: maximizing orthologous gene comparisons across deep evolutionary time. *PLoS One* 8.
- Henning, F., Lee, H.J., Franchini, P., et al., 2014. Genetic mapping of horizontal stripes in Lake Victoria cichlid fishes: benefits and pitfalls of using RAD markers for dense linkage mapping. *Mol. Ecol.* 23, 5224–5240.
- Hipp, A.L., Eaton, D.A., Cavender-Bares, J., et al., 2014. A framework phylogeny of the American oak clade based on sequenced RAD data. *PLoS One* 9, e93975.
- Hipp, A.L., Manos, P.S., González-Rodríguez, A., et al., 2018. Sympatric parallel diversification of major oak clades in the Americas and the origins of Mexican species diversity. *New Phytol.* 217, 439–452.
- Huang, H., Knowles, L.L., 2014. Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Syst. Biol.* 65, 357–365.
- Huang, D.I., Hefer, C.A., Kolosova, N., et al., 2014. Whole plastome sequencing reveals deep plastid divergence and cytonuclear discordance between closely related balsam poplars, *Populus balsamifera* and *P. trichocarpa* (Salicaceae). *New Phytol.* 204, 693–703.
- Huang, C.H., Sun, R., Hu, Y., et al., 2015. Resolution of Brassicaceae phylogeny using nuclear genes uncovers nested radiations and supports convergent morphological evolution. *Mol. Biol. Evol.* 33, 394–412.
- Huang, C.H., Zhang, C.F., Liu, M., et al., 2016. Multiple polyploidization events across Asteraceae with two nested events in the early history revealed by nuclear phylogenomics. *Mol. Biol. Evol.* 33, 2820–2835.
- Ilves, K.L., Lopez-Fernandez, H., 2014. A targeted next-generation sequencing toolkit for exon-based cichlid phylogenomics. *Mol. Ecol. Resour.* 14, 802–811.
- International Human Genome Sequencing Consortium, 2004. Finishing the euchromatic sequence of the human genome. *Nature* 431, 931–945.
- Jansen, R.K., Raubeson, L.A., Boore, J.L., et al., 2005. Methods for obtaining and analyzing whole chloroplast genome sequences. In: Zimmer, E.A., Roalson, E.H. (Eds.), *Molecular Evolution: Producing the Biochemical Data*, Part B, pp. 348–384.
- Jansen, R.K., Cai, Z., Raubeson, L.A., et al., 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19369–19374.
- Jarvis, E.D., Mirarab, S., Aberer, A.J., et al., 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 346, 1320–1331.
- Jones, M.R., Good, J.M., 2016. Targeted capture in evolutionary and ecological genomics. *Mol. Ecol.* 25, 185–202.
- Leaché, A.D., Banbury, B.L., Felsenstein, J., et al., 2015. Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. *Syst. Biol.* 64, 1032–1047.
- Lee, E.K., Cibrián-Jaramillo, A., Kolokotronis, S.O., et al., 2011. A functional phylogenomic view of the seed plants. *PLoS Genet.* 7.
- Lemmon, E.M., Lemmon, A.R., 2013. High-throughput genomic data in systematics and phylogenetics. *Annu. Rev. Ecol. Syst.* 44, 99–121.
- Liu, Y., Cox, C.J., Wang, W., et al., 2014. Mitochondrial phylogenomics of early land plants: mitigating the effects of saturation, compositional heterogeneity, and codon-usage bias. *Syst. Biol.* 63, 862–878.
- Lonsdale, D.M., Brears, T., Hodge, T.P., et al., 1988. The plant mitochondrial genome-homologous recombination as a mechanism for generating heterogeneity. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 319, 149–163.
- Ma, P.F., Zhang, Y.X., Zeng, C.X., et al., 2014. Chloroplast phylogenomic analyses resolve deep-level relationships of an intractable bamboo tribe Arundinarieae (poaceae). *Syst. Biol.* 63, 933–950.
- Male, P.J.G., Bardon, L., Besnard, G., et al., 2014. Genome skimming by shotgun sequencing helps resolve the phylogeny of a pantropical tree family. *Mol. Ecol. Resour.* 14, 966–975.
- Mandel, J.R., Dikow, R.B., Funk, V.A., et al., 2014. A target enrichment method for gathering phylogenetic information from hundreds of loci: an example from the Compositae. *Appl. Plant Sci.* 2.
- Martin, W., Deusch, O., Stawski, N., et al., 2005. Chloroplast genome phylogenetics: why we need independent approaches to plant molecular evolution. *Trends Plant Sci.* 10, 203–209.
- Massanti, R., Reznicek, A.A., Knowles, L.L., 2016. Utilizing RADseq data for phylogenetic analysis of challenging taxonomic groups: a case study in *Carex* sect. *Racemosae*. *Am. J. Bot.* 103, 337–347.
- Matasci, N., Hung, L.H., Yan, Z.X., et al., 2014. Data access for the 1,000 Plants (1KP) project. *GigaScience* 3.
- McCormack, J.E., Faircloth, B.C., Crawford, N.G., et al., 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. *Genome Res.* 22, 746–754.
- McCormack, J.E., Harvey, M.G., Faircloth, B.C., et al., 2013a. A phylogeny of birds based on over 1,500 loci collected by target enrichment and high-throughput sequencing. *PLoS One* 8.
- McCormack, J.E., Hird, S.M., Zellmer, A.J., et al., 2013b. Applications of next-generation sequencing to phylogeography and phylogenetics. *Mol. Phylogenet. Evol.* 66, 526–538.
- Mirarab, S., Bayzid, M.S., Boussau, B., et al., 2014. Statistical binning enables an accurate coalescent-based estimation of the avian tree. *Science* 346, 1250463.
- Misof, B., Liu, S., Meusemann, K., et al., 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science* 346, 763–767.
- Moore, M.J., Bell, C.D., Soltis, P.S., et al., 2007. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19363–19368.
- Moore, M.J., Soltis, P.S., Bell, C.D., et al., 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4623–4628.
- Nater, A., Burri, R., Kawakami, T., et al., 2015. Resolving evolutionary relationships in closely related species with whole-genome sequencing data. *Syst. Biol.* 64, 1000–1017.
- Nicholls, J.A., Pennington, R.T., Koenen, E.J.M., et al., 2015. Using targeted enrichment of nuclear genes to increase phylogenetic resolution in the neotropical rain forest genus *Inga* (Leguminosae: mimosoideae). *Front. Plant Sci.* 6.
- Nock, C.J., Waters, D.L.E., Edwards, M.A., et al., 2011. Chloroplast genome sequences from total DNA for plant identification. *J. Plant Biotechnol.* 9, 328–333.
- Ohyama, K., Fukuzawa, H., Kohchi, T., et al., 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* 322, 572–574.
- Olson, M., 2007. Enrichment of super-sized resequencing targets from the human genome. *Nat. Methods* 4, 891–892.
- Palmer, J.D., Herbon, L.A., 1988. Plant mitochondrial-DNA evolves rapidly in structure, but slowly in sequence. *J. Mol. Evol.* 28, 87–97.
- Peterson, B.K., Weber, J.N., Kay, E.H., et al., 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* 7, e37135.
- Pfender, W.F., Saha, M.C., Johnson, E.A., et al., 2011. Mapping with RAD (restriction-site associated DNA) markers to rapidly identify QTL for stem rust resistance in *Lolium perenne*. *Theor. Appl. Genet.* 122, 1467–1480.
- Philippe, H., Delsuc, F., Brinkmann, H., et al., 2005. Phylogenomics. *Annu. Rev. Ecol. Syst.* 36, 541–562.
- Philippe, H., Brinkmann, H., Lavrov, D.V., et al., 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. *PLoS Biol.* 9.
- Raubeson, L.A., Jansen, R.K., 2005. Chloroplast genomes of plants. In: Henry, R.J. (Ed.), *Plant Diversity and Evolution: Genotypic and Phenotypic Variation in Higher Plants*. CABI Publishing, Wallingford, pp. 45–68.
- Ren, G., Mateo, R.G., Liu, J., et al., 2017. Genetic consequences of Quaternary climatic oscillations in the Himalayas: *Primula tibetica* as a case study based on restriction site-associated DNA sequencing. *New Phytol.* 213, 1500–1512.
- Rokas, A., Holland, P.W., 2000. Rare genomic changes as a tool for phylogenetics. *Trends Ecol. Evol.* 15, 454–459.
- Ross, T.G., Barrett, C.F., Soto Gomez, M., et al., 2015. Plastid phylogenomics and molecular evolution of Alismatales. *Cladistics* 32, 160–178.
- Rubin, B.E.R., Ree, R.H., Moreau, C.S., 2012. Inferring phylogenies from RAD sequence data. *PLoS One* 7, e33394.
- Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463–5467.
- Schmickl, R., Liston, A., Zeisek, V., et al., 2016. Phylogenetic marker development for target enrichment from transcriptome and genome skim data: the pipeline and its application in southern African *Oxalis* (Oxalidaceae). *Mol. Ecol. Resour.* 16, 1124–1135.
- Schott, R.K., Panesar, B., Card, D.C., et al., 2017. Targeted capture of complete coding regions across divergent species. *Genome Biol. Evol.* 9, 398–414.
- Senapathy, P., Bhasi, A., Mattox, J., et al., 2010. Targeted genome-wide enrichment of functional regions. *PLoS One* 5, e11138.
- Shendure, J., Balasubramanian, S., Church, G.M., et al., 2017. DNA sequencing at 40: past, present and future. *Nature* 550, 345–353.
- Shinozaki, K., Ohme, M., Tanaka, M., et al., 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO J.* 5, 2043–2049.
- Sims, G.E., Jun, S.R., Wu, G.A., et al., 2009. Whole-genome phylogeny of mammals: evolutionary information in genic and nongenic regions. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17077–17082.
- Sollars, E.S., Harper, A.L., Kelly, L.J., et al., 2016. Genome sequence and genetic diversity of European ash trees. *Nature* 541, 212–216.
- Soltis, D.E., Gitzendanner, M.A., Stull, G., et al., 2013. The potential of genomics in plant systematics. *Taxon* 62, 886–898.
- Stamatakis, A., Aberer, A.J., Goll, C., et al., 2012. RAxML-Light: a tool for computing terabyte phylogenies. *Bioinformatics* 28, 2064–2066.
- Stegemann, S., Keuthe, M., Greiner, S., et al., 2012. Horizontal transfer of chloroplast genomes between plant species. *Proc. Natl. Acad. Sci. U.S.A.* 109, 2434–2438.
- Straub, S.C., Parks, M., Weitemier, K., et al., 2012. Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. *Am. J. Bot.* 99, 349–364.
- Sullivan, A.R., Schiffthaler, B., Thompson, S.L., et al., 2017. Interspecific plastome recombination reflects ancient reticulate evolution in *Picea* (pinaceae). *Mol. Biol. Evol.* 34, 1689–1701.
- Teh, B.T., Lim, K., Yong, C.H., et al., 2017. The draft genome of tropical fruit durian (*Durio zibethinus*). *Nat. Genet.* 49, 1633–1641.
- The *C. elegans* Sequencing Consortium, 1998. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282, 2012–2018.
- Tsitrone, A., Kirkpatrick, M., Levin, D.A., 2003. A model for chloroplast capture. *Evolution* 57, 1776–1782.

- Valderrama, E., Richardson, J.E., Kidner, C.A., et al., 2018. Transcriptome mining for phylogenetic markers in a recently radiated genus of tropical plants (*Renealmia* L.f., Zingiberaceae). *Mol. Phylogenet. Evol.* 119, 13–24.
- Vargas, O.M., Ortiz, E.M., Simpson, B.B., 2017. Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: astereae: *Diplostephium*). *New Phytol.* 214, 1736–1750.
- Wagner, C.E., Keller, I., Wittwer, S., et al., 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Mol. Ecol.* 22, 787–798.
- Wang, X., Ye, X., Zhao, L., et al., 2017. Genome-wide RAD sequencing data provide unprecedented resolution of the phylogeny of temperate bamboos (Poaceae: bambusoideae). *Sci. Rep.* 7, 11546.
- Weitemier, K., Straub, S.C., Cronn, R.C., et al., 2014. Hyb-Seq: combining target enrichment and genome skimming for plant phylogenomics. *Appl. Plant Sci.* 2, 1400042.
- Wickett, N.J., Mirarab, S., Nguyen, N., et al., 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proc. Natl. Acad. Sci. U. S. A.* 111, E4859–E4868.
- Wysocki, W.P., Clark, L.G., Attigala, L., et al., 2015. Evolution of the bamboos (Bambusoideae; Poaceae): a full plastome phylogenomic analysis. *BMC Evol. Biol.* 15.
- Xiang, Y.Z., Huang, C.H., Hu, Y., et al., 2016. Evolution of Rosaceae fruit types based on nuclear phylogeny in the context of geological times and genome duplication. *Mol. Biol. Evol.* 34, 262–281.
- Yang, J.B., Li, D.Z., Li, H.T., 2014. Highly effective sequencing whole chloroplast genomes of Angiosperms by nine novel universal primer pairs. *Mol. Ecol. Resour.* 14, 1024–1031.
- Yang, Y., Moore, M.J., Brockington, S.F., et al., 2015. Dissecting molecular evolution in the highly diverse plant clade Caryophyllales using transcriptome sequencing. *Mol. Biol. Evol.* msv081.
- Yang, G.Q., Chen, Y.M., Wang, J.P., et al., 2016. Development of a universal and simplified ddRAD library preparation approach for SNP discovery and genotyping in angiosperm plants. *Plant Meth.* 12, 39.
- Yi, T.S., Jin, G.H., Wen, J., 2015. Chloroplast capture and intra- and inter-continental biogeographic diversification in the Asian – new World disjunct plant genus *Osmorhiza* (Apiaceae). *Mol. Phylogenet. Evol.* 85, 10–21.
- Yu, X.Q., Gao, L.M., Soltis, D.E., et al., 2017. Insights into the historical assembly of East Asian subtropical evergreen broadleaved forests revealed by the temporal history of the tea family. *New Phytol.* 215, 1235–1248.
- Zeng, L.P., Zhang, Q., Sun, R.R., et al., 2014. Resolution of deep Angiosperm phylogeny using conserved nuclear genes and estimates of early divergence times. *Nat. Commun.* 5, 4956.
- Zeng, L., Ming, C., Li, Y., et al., 2017a. Out of southern East Asia of the brown rat revealed by large scale genome sequencing. *Mol. Biol. Evol.* 35, 149–158.
- Zeng, L.P., Zhang, N., Zhang, Q., et al., 2017b. Resolution of deep eudicot phylogeny and their temporal diversification using nuclear genes from transcriptomic and genomic datasets. *New Phytol.* 214, 1338–1354.
- Zhang, N., Zeng, L.P., Shan, H.Y., et al., 2012. Highly conserved low-copy nuclear genes as effective markers for phylogenetic analyses in angiosperms. *New Phytol.* 195, 923–937.
- Zhang, T., Zeng, C.X., Yang, J.B., et al., 2016. Fifteen novel universal primer pairs for sequencing whole chloroplast genomes and a primer pair for nuclear ribosomal DNAs. *J. Systemat. Evol.* 54, 219–227.
- Zhang, G.Q., Liu, K.W., Li, Z., et al., 2017a. The *Apostasia* genome and the evolution of orchids. *Nature* 549, 379–383.
- Zhang, S.D., Jin, J.J., Chen, S.Y., et al., 2017b. Diversification of Rosaceae since the late cretaceous based on plastid phylogenomics. *New Phytol.* 214, 1355–1367.
- Zimmer, E.A., 2013. Reprint of: using nuclear gene data for plant phylogenetics: progress and prospects. *Mol. Phylogenet. Evol.* 66, 539–550.